

STIC-Biotech/ChemLib

9.6369

From: Ramirez, Delia  
Sent: Wednesday, June 11, 2003 12:43 PM  
To: STIC-Biotech/ChemLib  
Subject: case 09/911,860

Hi,

I would like to request the following search (commercial and interference):

1. a standard search of seq id 3 in the nucleic acid databases.

Thank you,

Delia M. Ramirez, Ph.D.  
Patent Examiner  
Recombinant Enzymes-Art Unit 1652  
USPTO  
1911 S. Clark Street, Crystal Mall 1, 10D06, Mail room 10D01  
Arlington, VA 22202  
(703) 306-0288  
delia.ramirez@uspto.gov

CRFE

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 6/13/03  
Date Completed: 6/11/03  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH: 1  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: 91  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 07:32:37 ; Search time 238 Seconds

(without alignments)  
170,319 Million cell updates/sec

Title: US-09-911-860A-3

Perfect score: 18

Sequence: 1 gcttcgggaactgaag 18

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : N.Geneseq\_101002.\*

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*

2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*

3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*

4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*

5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*

6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*

7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*

8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*

9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*

10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*

11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*

12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*

13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*

14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*

15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*

16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*

17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*

18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*

19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*

20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*

21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*

22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*

23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*

24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	100.0	18	24	ABK14145	Chlorinated ethylene
2	18	100.0	1212	22	AA62239	A 16S rDNA sequenc
3	18	100.0	1212	22	AA62241	A 16S rDNA sequenc
4	18	100.0	1212	22	AA62242	A 16S rDNA sequenc
5	18	100.0	1212	22	AA62243	A 16S rDNA sequenc
6	18	100.0	1335	22	AA62240	A 16S rDNA sequenc
7	18	100.0	1443	22	AA62244	A 16S rDNA sequenc
8	15.4	85.6	244	24	ABN87653	Human prostate spe
9	15.4	85.6	531	24	ABN87664	Human prostate spe

10	15.4	85.6	560	22	ABA63560	Human foetal liver
11	15.4	85.6	560	22	ABA76630	Human foetal liver
12	15.4	85.6	560	22	ABA30755	Probe #9221 for ge
13	15.4	85.6	560	22	ABA41138	Probe #19604 for g
14	15.4	85.6	560	22	AAK12088	Human brain expres
15	15.4	85.6	560	22	AAK25265	Human brain expres
16	15.4	85.6	560	22	AAK37803	Human bone marrow
17	15.4	85.6	560	22	AAK51270	Human bone marrow
18	15.4	85.6	560	22	AAI18558	Probe #8491 for ge
19	15.4	85.6	560	22	AAI28281	Probe #18214 for g
20	15.4	85.6	560	22	AAI43673	Probe #12359 used
21	15.4	85.6	560	22	AAI57326	Probe #26012 used
22	15.4	85.6	560	24	ABSI1794	Human genome-deriv
23	15.4	85.6	560	24	ABS24809	Mammalian AMPK alp
24	15.4	85.6	1647	18	AA85925	Knockout mouse dis
25	15.4	85.6	1647	18	AA85925	Drosophila melanog
26	15.4	85.6	2757	23	ABL23365	Drosophila melanog
27	15.4	85.6	6313	23	ABL23364	Drosophila melanog
28	15.4	85.6	6773	21	AAAO9326	Human cancer assoc
29	15.4	85.6	6773	21	AAAO9799	Human ovarian tumo
30	15.4	85.6	10669	22	AAK76914	Human immune/haema
31	15.4	85.6	10669	22	AAK81696	Human immune/haema
32	15.4	85.6	15441	23	ABLO6029	Drosophila melanog
33	15.4	85.6	16951	22	AAH48620	Human fascin DNA f
34	15.4	85.6	16951	22	AAH48622	Human fascin DNA f
35	15.4	85.6	20706	23	ABLO6028	Drosophila melanog
36	15.4	85.6	23432	22	AA535832	Human cardiovascular
37	15	83.3	2340	24	ABK35812	CDNA sequence #203
38	15	83.3	9021	20	AAK13092	Enterococcus faeca
39	15	83.3	40352	19	AAV02032	MAGE-B cluster DNA
40	14.8	82.2	420	17	AAV13658	ACNPV ORF 53, resi
41	14.8	82.2	540	19	AAV15410	Vancomycin resista
42	14.8	82.2	555	22	AAK88737	Human digestive sy
43	14.8	82.2	555	22	AA531771	Human liver associ
44	14.8	82.2	535	24	ABN90126	Human liver antige
45	14.8	82.2	648	23	ABLO3049	Drosophila melanog

## ALIGNMENTS

RESULT 1  
ABK14145  
ID ABK14145 standard; DNA; 18 BP.

XX  
AC ABK14145:

XX  
DT 08-MAY-2002 (first entry)

XX  
DE Chlorinated ethylene-decomposing bacteria detection DNA Kwi-De3.

KW Chlorinated ethylene-decomposing bacteria; 16S rRNA; 16S rRNA; ss; probe.

KW PCR; primer; sol; underground water; chlorinated ethylene; Kwi-De3;

KW chlorinated ethane; Dehalococoides.

XX  
OS Synthetic.

XX  
PN EP1176216-A2.

XX  
PD 30-JAN-2002.

XX  
PF 23-JUL-2001; 2001EP-0117844.

XX  
PR 24-JUL-2000; 2000JP-0227580.

XX  
PR 09-MAR-2001; 2001JP-0066001.

XX  
PA (KURK ) KURITA WATER IND LTD.

XX  
PI Nakamura K, Ueno T;

XX  
DR WPI; 2002-173127/23.

XX  
PT New nucleic acid for detecting chlorinated ethylene-decomposing

PT bacteria used to purify soil or underground water contaminated with  
 PT chlorinated ethylene or ethane  
 PS Claim 1; Page 7; 22pp; English.  
 XX  
 CC The invention relates to a nucleic acid which hybridises to the 16S  
 CC ribosomal (deoxy)ribonucleic acid of chlorinated ethylene-decomposing  
 CC bacteria. The nucleic acid can be used as a labelled probe for detecting  
 CC chlorinated ethylene-decomposing bacteria (e.g. Dehalococcoides)  
 CC comprising the novel nucleic acid by DNA hybridisation using the labelled  
 CC probe as an indicator. The bacteria can also be detected by performing  
 CC PCR using the nucleic acid as a primer and the sample nucleic acid as a  
 CC template, and detecting newly synthesised DNA. A method for decomposing  
 CC chlorinated ethylene or ethane comprises detecting chlorinated  
 CC ethylene-decomposing bacteria using underground water or soil as a  
 CC sample, and introducing the water/soil containing the bacteria to soil  
 CC or underground water contaminated by chlorinated ethylene or ethane. The  
 CC methods are therefore useful for purifying soil or underground water  
 CC contaminated with chlorinated ethylene or ethane. This sequence  
 CC represents a nucleic acid which hybridises to nucleic acid of chlorinated  
 CC ethylene-decomposing bacteria.  
 CC  
 SQ Sequence 18 BP; 5 A; 3 C; 7 G; 3 T; 0 other;  
 Query Match 100.0%; Score 18; DB 24; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 5.6;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCTTCGGGAACCTGAAGC 18  
 Db 1 GCTTCGGGAACCTGAAGC 18  
 RESULT 2  
 AAC62239  
 ID AAC62239 standard; DNA; 1212 BP.  
 AC AAC62239;  
 XX  
 DT 19-MAR-2001 (first entry)  
 XX  
 DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.  
 XX  
 KM 16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride;  
 KM carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;  
 KM trichloroethane; dichloroethylene; chlorinating bacteria; ss.  
 XX  
 OS Dehalococcoides ethenogenes.  
 XX  
 PN WO200063443-A2.  
 XX  
 PD 26-OCT-2000.  
 XX  
 PF 13-APR-2000; 2000MO-US09883.  
 XX  
 PR 15-APR-1999; 99US-0129511.  
 XX  
 PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX  
 PI Hendrickson ER, Ebersole RC;  
 XX  
 DR WPI; 2001-024581/03.  
 XX  
 PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and  
 PT indicative of a dechlorinating bacterial strain -  
 XX  
 PS Claim 1; Page 47; 55pp; English.  
 XX  
 CC The present sequence represents the 16S rDNA profile of Dehalococcoides  
 CC ethenogenes strain PL, isolated from soil surrounding an industrial  
 CC site. The 16S rDNA profile is linked to dechlorinating activity.  
 CC Bacterial strain comprising the 16S rDNA sequence of the invention are  
 CC useful for the dechlorination of chlorinated compounds such as

CC carbontetrachloride, tetrachloroethane, chloroform, dichloromethane,  
 CC trichloroethane, dichloroethylene, vinyl chloride, and chloroformatics.  
 CC The 16S rDNA sequence is also useful for identification of new  
 CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides  
 CC ethenogenes.  
 CC  
 SQ Sequence 1212 BP; 321 A; 252 C; 369 G; 269 T; 1 other;  
 Query Match 100.0%; Score 18; DB 22; Length 1212;  
 Best Local Similarity 100.0%; Pred. No. 9.5;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCTTCGGGAACCTGAAGC 18  
 Db 112 GCTTCGGGAACCTGAAGC 129  
 RESULT 3  
 AAC62241  
 ID AAC62241 standard; DNA; 1212 BP.  
 AC AAC62241;  
 XX  
 DT 19-MAR-2001 (first entry)  
 XX  
 DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.  
 XX  
 KM 16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride;  
 KM carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;  
 KM trichloroethane; dichloroethylene; chlorinating bacteria; ss.  
 XX  
 OS Dehalococcoides ethenogenes.  
 XX  
 PN WO200063443-A2.  
 XX  
 PD 26-OCT-2000.  
 XX  
 PF 13-APR-2000; 2000MO-US09883.  
 XX  
 PR 15-APR-1999; 99US-0129511.  
 XX  
 PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX  
 PI Hendrickson ER, Ebersole RC;  
 XX  
 DR WPI; 2001-024581/03.  
 XX  
 PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and  
 PT indicative of a dechlorinating bacterial strain -  
 XX  
 PS Claim 1; Page 48; 55pp; English.  
 XX  
 CC The present sequence represents the 16S rDNA profile of Dehalococcoides  
 CC ethenogenes strain DAB, isolated from soil surrounding an industrial  
 CC site. The 16S rDNA profile is linked to dechlorinating activity.  
 CC Bacterial strain comprising the 16S rDNA sequence of the invention are  
 CC useful for the dechlorination of chlorinated compounds such as  
 CC carbontetrachloride, tetrachloroethane, chloroform, dichloromethane,  
 CC trichloroethane, dichloroethylene, vinyl chloride, and chloroformatics.  
 CC The 16S rDNA sequence is also useful for identification of new  
 CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides  
 CC ethenogenes.  
 CC  
 SQ Sequence 1212 BP; 318 A; 253 C; 372 G; 269 T; 0 other;  
 Query Match 100.0%; Score 18; DB 22; Length 1212;  
 Best Local Similarity 100.0%; Pred. No. 9.5;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCTTCGGGAACCTGAAGC 18  
 Db 112 GCTTCGGGAACCTGAAGC 129

```

RESULT 4
AAC62242
ID AAC62242 standard; DNA: 1212 BP.
XX
AC AAC62242;
XX
DT 19-MAR-2001 (first entry)
XX
DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.
XX
KM 16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride;
KW carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;
KM trichloroethane; dichloroethylene; chlorinating bacteria; ss.
XX
OS Dehalococcoides ethenogenes.
XX
PN WO200063443-A2.
XX
PD 26-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US09883.
XX
PR 15-APR-1999; 99US-0129511.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX
PI Hendrickson ER, Ebersole RC;
XX
DR WPI; 2001-024581/03.
XX
PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and
XX indicative of a dechlorinating bacterial strain -
XX
PS Claim 1; Page 48-49; 55pp; English.
XX
CC The present sequence represents the 16S rDNA profile of Dehalococcoides
CC ethenogenes strain PIN, isolated from soil surrounding an industrial
CC site. The 16S rDNA profile is linked to dechlorinating activity.
CC Bacterial strain comprising the 16S rDNA sequence of the invention are
CC useful for the dechlorination of chlorinated compounds such as
CC carbontetrachloride, tetrachloroethane, chloroform, dichloromethane,
CC trichloroethane, dichloroethylene, vinyl chloride, and chloroformatics.
CC The 16S rDNA sequence is also useful for identification of new
CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides
CC ethenogenes.
XX
SO Sequence 1212 BP; 320 A; 253 C; 371 G; 268 T; 0 other;

Query Match 100.0%; Score 18; DB 22; Length 1212;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTCGGGAAACTGAAGG 18
   ||||||||||||||||
DB 112 GCTTCGGGAAACTGAAGG 129

RESULT 5
AAC62243
ID AAC62243 standard; DNA: 1212 BP.
XX
AC AAC62243;
XX
DT 19-MAR-2001 (first entry)
XX
DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.
XX
KM 16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride;
KW carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;
KM trichloroethane; dichloroethylene; chlorinating bacteria; ss.
XX
OS Dehalococcoides ethenogenes.

```

```

XX
PN WO200063443-A2.
XX
PD 26-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US09883.
XX
PR 15-APR-1999; 99US-0129511.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX
PI Hendrickson ER, Ebersole RC;
XX
DR WPI; 2001-024581/03.
XX
PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and
XX indicative of a dechlorinating bacterial strain -
XX
PS Claim 1; Page 49; 55pp; English.
XX
CC The present sequence represents the 16S rDNA profile of Dehalococcoides
CC ethenogenes strain DLU, isolated from soil surrounding an industrial
CC site. The 16S rDNA profile is linked to dechlorinating activity.
CC Bacterial strain comprising the 16S rDNA sequence of the invention are
CC useful for the dechlorination of chlorinated compounds such as
CC carbontetrachloride, tetrachloroethane, chloroform, dichloromethane,
CC trichloroethane, dichloroethylene, vinyl chloride, and chloroformatics.
CC The 16S rDNA sequence is also useful for identification of new
CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides
CC ethenogenes.
XX
SO Sequence 1212 BP; 320 A; 255 C; 371 G; 266 T; 0 other;

Query Match 100.0%; Score 18; DB 22; Length 1212;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTCGGGAAACTGAAGG 18
   ||||||||||||||||
DB 112 GCTTCGGGAAACTGAAGG 129

RESULT 6
AAC62240
ID AAC62240 standard; DNA: 1335 BP.
XX
AC AAC62240;
XX
DT 19-MAR-2001 (first entry)
XX
DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.
XX
KM 16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride;
KW carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;
KM trichloroethane; dichloroethylene; chlorinating bacteria; ss.
XX
OS Dehalococcoides ethenogenes.
XX
PN WO200063443-A2.
XX
PD 26-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US09883.
XX
PR 15-APR-1999; 99US-0129511.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX
PI Hendrickson ER, Ebersole RC;
XX
DR WPI; 2001-024581/03.
XX
PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and

```

PT Indicative of a dechlorinating bacterial strain -  
XX  
PS Claim 1; Page 47-48; 55pp; English.  
XX  
CC The present sequence represents the 16S rDNA profile of Dehalococcoides  
CC ethenogenes strain 195, isolated from soil surrounding an industrial  
CC site. The 16S rDNA profile is linked to dechlorinating activity.  
CC Bacterial strain comprising the 16S rDNA sequence of the invention are  
CC useful for the dechlorination of chlorinated compounds such as  
CC carbontetrachloride, tetrachloroethane, chloroform, dichloromethane,  
CC trichloroethane, dichloroethylene, vinyl chloride, and chloroacromatics.  
CC The 16S rDNA sequence is also useful for identification of new  
CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides  
CC ethenogenes.  
XX  
SQ Sequence 1335 BP; 348 A; 288 C; 403 G; 296 T; 0 other:  
Query Match 100.0%; Score 18; DB 22; Length 1335;  
Best Local Similarity 100.0%; Pred. No. 9.6;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GCTTCGGGAAACTGAGG 18  
DB 112 GCTTCGGGAAACTGAGG 129  
RESULT 7  
AAC62244  
ID AAC62244 standard; DNA; 1443 BP.  
XX  
AC AAC62244;  
XX  
DT 19-MAR-2001 (first entry)  
XX  
DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.  
XX  
XX 16S rDNA: dechlorinating activity; chlorinated compound; vinyl chloride;  
KW carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;  
KW trichloroethane; dichloroethylene; chlorinating bacteria; ss.  
XX  
OS Dehalococcoides ethenogenes.  
XX  
PN MO200063443-A2.  
XX  
PD 26-OCT-2000.  
XX  
PF 13-APR-2000; 2000MO-US09883.  
XX  
PR 15-APR-1999; 99US-0129511.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E. I.  
XX  
PI Hendrickson ER, Ebersole RC;  
XX  
DR WPI: 2001-024581/03.  
XX  
PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and  
PT indicative of a dechlorinating bacterial strain -  
XX  
PS Example 2; Page 49-50; 55pp; English.  
XX  
CC The present sequence represents the 16S rDNA profile of Dehalococcoides  
CC ethenogenes strain 195. The 16S rDNA profile is linked to dechlorinating  
CC activity. Bacterial strain comprising the 16S rDNA sequence of the  
CC invention are useful for the dechlorination of chlorinated compounds such  
CC as carbontetrachloride, tetrachloroethane, chloroform, dichloromethane,  
CC trichloroethane, dichloroethylene, vinyl chloride, and chloroacromatics.  
CC The 16S rDNA sequence is also useful for identification of new  
CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides  
CC ethenogenes.  
XX  
SQ Sequence 1443 BP; 379 A; 306 C; 443 G; 314 T; 1 other:

Query Match 100.0%; Score 18; DB 22; Length 1443;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GCTTCGGGAAACTGAGG 18  
DB 112 GCTTCGGGAAACTGAGG 129  
RESULT 8  
ABN87653/c  
ID ABN87653 standard; CDNA; 244 BP.  
XX  
AC ABN87653;  
XX  
DT 08-AUG-2002 (first entry)  
XX  
DE Human prostate specific gene cDNA sequence SEQ ID NO:14.  
XX  
KW Human; prostate specific gene; prostate specific protein; PSG; PSP;  
KW prostate cancer; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO200236808-A2.  
XX  
PD 10-MAY-2002.  
XX  
PF 05-NOV-2001; 2001MO-US47283.  
XX  
PR 03-NOV-2000; 2000US-245740P.  
XX  
PA (DIAD-) DIADEXUS INC.  
XX  
PI Sun Y, Reclpon H, Chen S, Liu C;  
XX  
DR WPI: 2002-471506/50.  
XX  
PT New prostate-specific nucleic acids and polypeptides, useful for  
PT identifying, diagnosing, monitoring, staging, imaging, and treating  
PT prostate cancer and non-cancerous disease states in prostate tissue  
XX  
PS Claim 1; Page 165; 254pp; English.  
XX  
CC ABN87650 to ABN87789 represent human prostate-specific nucleic acids (I),  
CC and ABN79192 to ABN79295 represent human prostate-specific proteins (II)  
CC from the present invention. (I) and (II) have cytostatic activity. (I)  
CC can be used in gene therapy. The prostate-specific nucleic acids,  
CC polypeptides and compositions from the present invention can be used for  
CC identifying, diagnosing, monitoring, staging, imaging, and treating  
CC prostate cancer and non-cancerous disease states in prostate tissue; for  
CC identifying prostate tissue; for monitoring, identifying and/or designing  
CC agonists and antagonists of the polypeptides; in gene therapy; in  
CC producing transgenic animals and cells; for producing engineered prostate  
CC tissue for treatment and research; and as elements in an array or  
CC computer program for pattern recognition of prostate disorders. The  
CC nucleic acids may be used as hybridisation probes to detect, characterise  
CC and quantify hybridising nucleic acids in, and isolate hybridising  
CC nucleic acids from, both genomic and transcript-derived nucleic acid  
CC samples.  
XX  
SQ Sequence 244 BP; 56 A; 73 C; 64 G; 51 T; 0 other:  
XX  
Query Match 85.6%; Score 15.4; DB 24; Length 244;  
Best Local Similarity 94.1%; Pred. No. 1.8e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 GCTTCGGGAAACTGAGG 17  
DB 27 GCTTCGGGAAACTGAGG 11  
RESULT 9

ABN87664  
ID ABN87664 standard; cDNA; 531 BP.  
XX  
AC ABN87664;  
XX  
DT 08-AUG-2002 (first entry)  
XX  
DE Human prostate specific gene cDNA sequence SEQ ID NO:15.  
XX  
KW Human; prostate specific gene; prostate specific protein; PSG; PSP;  
XX prostate cancer; chromosome 8; gene; ss.  
OS Homo sapiens.  
XX  
PN WO200236808-A2.  
XX  
PD 10-MAY-2002.  
XX  
PE 05-NOV-2001; 2001WO-US47283.  
XX  
PR 03-NOV-2000; 2000US-245740P.  
XX  
PA (DIAD-) DIADEXUS INC.  
XX  
PI Sun Y, Reclipon H, Chen S, Liu C;  
XX  
DR WPI; 2002-471506/50.  
XX  
PT New prostate-specific nucleic acids and polypeptides, useful for  
PT identifying, diagnosing, monitoring, staging, imaging, and treating  
PT prostate cancer and non-cancerous disease states in prostate tissue  
XX  
PS Claim 1; Page 165; 254pp; English.  
XX  
CC ABN87650 to ABN87789 represent human prostate-specific nucleic acids (I),  
CC and ABN79192 to ABN79295 represent human prostate-specific proteins (II)  
CC from the present invention. (I) and (II) have cytostatic activity. (I)  
CC can be used in gene therapy. The prostate-specific nucleic acids,  
CC polypeptides and compositions from the present invention can be used for  
CC identifying, diagnosing, monitoring, staging, imaging, and treating  
CC prostate cancer and non-cancerous disease states in prostate tissue; for  
CC identifying prostate tissue; for monitoring, identifying and/or designing  
CC agonists and antagonists of the polypeptides; in gene therapy; in  
CC producing transgenic animals and cells; for producing engineered prostate  
CC tissue for treatment and research; and as elements in an array or  
CC computer program for pattern recognition of prostate disorders. The  
CC nucleic acids may be used as hybridisation probes to detect, characterise  
CC and quantify hybridising nucleic acids in, and isolate hybridising  
CC nucleic acids from, both genomic and transcript-derived nucleic acid  
CC samples.  
XX  
SQ Sequence 531 BP; 102 A; 147 C; 143 G; 135 T; 4 other;  
XX  
Query Match 85.6%; Score 15.4; DB 24; Length 531;  
Best Local Similarity 94.1%; Pred. No. 2e+02; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCTTCGGGAAACTGAG 17  
DB 505 GCTTGGGAAACTGAG 521  
|||||  
RESULT 10  
ABA63560  
ID ABA63560 standard; DNA; 560 BP.  
XX  
AC ABA63560;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Human foetal liver single exon nucleic acid probe #11865.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SC, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human fetal liver -  
XX  
PS Claim 1; SEQ ID NO 11865; 639pp + sequence listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
XX  
SQ Sequence 560 BP; 132 A; 134 C; 126 G; 168 T; 0 other;  
XX  
Query Match 85.6%; Score 15.4; DB 22; Length 560;  
Best Local Similarity 94.1%; Pred. No. 2e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCTTCGGGAAACTGAG 17  
DB 64 GCTTCGGGAAACTGAG 80  
|||||  
RESULT 11  
ABA76630  
ID ABA76630 standard; DNA; 560 BP.  
XX  
AC ABA76630;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Human foetal liver single exon nucleic acid probe #24935.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.



CC from the human heart via microarrays. By measuring gene expression, the  
CC probes are useful for predicting, diagnosing, grading, staging,  
CC monitoring and prognosing diseases of the human heart and vascular system  
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
CC congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp://ipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 560 BP; 132 A; 134 C; 126 G; 168 T; 0 other;

Query Match 85.6%; Score 15.4; DB 22; Length 560;  
Best Local Similarity 94.1%; Pred. No. 2e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACCTGAAG 17  
||||| |||||||||  
DB 64 GCTTCGGGAACCTGAAG 80

RESULT 14

AAK12088 ID AAK12088 standard; DNA; 560 BP.

XX AAK12088;

XX 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 12079.

XX Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer; ss.

XX Homo sapiens.

OS WO200157275-A2.

PN 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

XX Example 4; SEQ ID NO: 12079; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is one of the probes of the

XX Sequence 560 BP; 132 A; 134 C; 126 G; 168 T; 0 other;

Query Match 85.6%; Score 15.4; DB 22; Length 560;  
Best Local Similarity 94.1%; Pred. No. 2e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACCTGAAG 17  
||||| |||||||||  
DB 64 GCTTCGGGAACCTGAAG 80

RESULT 15

AAK25265 ID AAK25265 standard; DNA; 560 BP.

XX AAK25265;

XX 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 25256.

XX Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer; ss.

XX Homo sapiens.

OS WO200157275-A2.

PN 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

XX Example 4; SEQ ID NO: 25256; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is one of the probes of the

XX Sequence 560 BP; 132 A; 134 C; 126 G; 168 T; 0 other;

Query Match 85.6%; Score 15.4; DB 22; Length 560;  
Best Local Similarity 94.1%; Pred. No. 2e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACCTGAAG 17  
||||| |||||||||  
DB 64 GCTTCGGGAACCTGAAG 80

Search completed: June 19, 2003, 08:48:43  
Job time : 245 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 07:21:02 ; Search time 147 Seconds  
(without alignments)  
179.685 Million cell updates/sec

Title: US-09-911-860a-3

Sequence: 1 gcttcgggaactgaagg 18

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:\*

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	100.0	18	9	US-09-911-860a-3	Sequence 3, Appl1
2	18	100.0	1377	9	US-10-061-071-2	Sequence 2, Appl1
3	18	100.0	1377	9	US-10-061-071-4	Sequence 4, Appl1
4	18	100.0	1377	9	US-10-061-071-5	Sequence 5, Appl1
5	18	100.0	1377	9	US-10-061-071-6	Sequence 6, Appl1
6	18	100.0	1378	9	US-10-061-071-3	Sequence 3, Appl1
7	18	100.0	1443	9	US-10-061-071-7	Sequence 7, Appl1
8	16.4	91.1	207	10	US-09-878-574-6798	Sequence 6798, Ap
9	15.4	85.6	244	9	US-10-011-585a-14	Sequence 14, Appl1
10	15.4	85.6	459	9	US-09-918-995-36193	Sequence 36193, A
11	15.4	85.6	473	9	US-10-102-524-314	Sequence 314, App
12	15.4	85.6	507	10	US-09-998-598-663	Sequence 663, App
13	15.4	85.6	531	9	US-10-011-585a-15	Sequence 15, Appl1
14	15.4	85.6	560	10	US-09-864-761-9221	Sequence 9221, Ap
15	15.4	85.6	560	10	US-09-864-761-26458	Sequence 26458, A
16	15.4	85.6	6773	10	US-09-864-864-336	Sequence 336, App
17	15.4	85.6	23432	10	US-10-091-504-1332	Sequence 1332, Ap
18	15.4	85.6	23432	10	US-09-764-869-1332	Sequence 1332, Ap
19	15	83.3	2340	9	US-09-822-846-203	Sequence 203, App

20	15	83.3	9021	10	US-09-070-927a-155	Sequence 155, App
21	14.8	82.2	555	9	US-10-073-961-101	Sequence 101, App
22	14.8	82.2	555	10	US-09-764-887-101	Sequence 101, App
23	14.8	82.2	1929	9	US-09-938-842a-1109	Sequence 1109, App
24	14.8	82.2	3668	10	US-09-759-667a-2	Sequence 2, Appl1
25	14.8	82.2	4649	10	US-10-198-846-10393	Sequence 10393, A
26	14.8	82.2	10156	10	US-09-759-667a-1	Sequence 1, Appl1
27	14.8	82.2	1503841	9	US-09-946-807-1	Sequence 1, Appl1
28	14.8	82.2	1503841	10	US-09-795-668-1	Sequence 1, Appl1
29	14.8	82.2	1503841	10	US-09-795-668-1	Sequence 1, Appl1
30	14.4	80.0	384	9	US-09-918-995-5841	Sequence 5841, Ap
31	14.4	80.0	401	9	US-09-918-995-33990	Sequence 33990, A
32	14.4	80.0	470	9	US-09-918-995-829	Sequence 829, App
33	14.4	80.0	510	9	US-09-918-995-32231	Sequence 32231, A
34	14.4	80.0	658	9	US-10-198-846-8463	Sequence 8463, Ap
35	14.4	80.0	874	9	US-10-198-846-4799	Sequence 4799, Ap
36	14.4	80.0	876	9	US-10-198-846-12484	Sequence 12484, A
37	14.4	80.0	2000	9	US-09-938-842a-4533	Sequence 4533, Ap
38	14.4	80.0	9792	10	US-09-895-072-14	Sequence 14, Appl1
39	14.4	80.0	9792	10	US-09-986-552-14	Sequence 14, Appl1
40	14.4	80.0	41907	10	US-09-967-013-5	Sequence 5, Appl1
41	14.4	80.0	42521	9	US-09-373-658-23	Sequence 23, Appl1
42	14	77.8	25	9	US-10-098-263b-72849	Sequence 72849, A
43	14	77.8	279	10	US-09-878-574-14116	Sequence 14116, A
44	14	77.8	451	9	US-10-060-036-2049	Sequence 2049, Ap
45	14	77.8	581	10	US-09-864-761-12485	Sequence 12485, A

## ALIGNMENTS

```
RESULT 1
US-09-911-860a-3
: Sequence 3, Application US/09911860A
: Publication No. US20030104383A1
: GENERAL INFORMATION:
: APPLICANT: Ueno, Toshihiro
: TITLE OF INVENTION: Nucleic Acid, Nucleic Acid for Detecting Chlorinated Ethylene-
: TITLE OF INVENTION: Bacteria, Probe, Method of Detecting Chlorinated Ethylene-Dec
: FILE REFERENCE: 9659/01377-US0
: CURRENT APPLICATION NUMBER: US/09/911, 860A
: CURRENT FILING DATE: 2002-12-17
: PRIOR APPLICATION NUMBER: JP2000-227580
: PRIOR FILING DATE: 2000-07-24
: PRIOR APPLICATION NUMBER: JP2001-066001
: PRIOR FILING DATE: 2001-03-09
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 3
: LENGTH: 18
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: primer
US-09-911-860a-3

Query Match      100.0%: Score 18; DB 9; Length 18;
Best Local Similarity 100.0%: Pred. No. 1.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GCTTCGGGAACGTGAAG 18
Db      1 GCTTCGGGAACGTGAAG 18

RESULT 2
US-10-061-071-2
: Sequence 2, Application US/10061071
: Publication No. US20030077601A1
: GENERAL INFORMATION:
: APPLICANT: E.I. DUPONT DE NEMOURS & COMPANY
```

RESULT 4  
US-10-061-071-5  
; Sequence 5, Application US/10061071  
; Publication NO. US20030077601A1  
; GENERAL INFORMATION:  
; APPLICANT: E.I. DUPONT DE NEMOURS & COMPANY  
; APPLICANT: HENDRICKSON, EDWIN  
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS FOR THE IDENTIFICATION OF DECHLORINATING  
; TITLE OF INVENTION: BACTERIA  
; FILE REFERENCE: BC1002 US CIP  
; CURRENT APPLICATION NUMBER: US/10/061,071  
; CURRENT FILING DATE: 2002-01-29  
; PRIOR APPLICATION NUMBER: US 60/129,511

```

1 RESULT 6
2 US-10-061-071-3
3 ; Sequence 3, Application US/10061071
4 ; Publication No. US20030077601A1
5 ; GENERAL INFORMATION:
6 ; APPLICANT: E.I. DUPONT DE NEMOURS & COMPANY
7 ; APPLICANT: HENDRICKSON, EDWIN
8 ; TITLE OF INVENTION: NOCETIC ACID FRAGMENTS FOR THE IDENTIFICATION OF DECHLORINATIN
9 ; TITLE OF INVENTION: BACTERIA
10 ; FILE REFERENCE: BC1002 US CIP
11 ; CURRENT APPLICATION NUMBER: US/10/061,071
12 ; CURRENT FILING DATE: 2002-01-29
13 ; PRIOR APPLICATION NUMBER: US 60/129,511
14 ; PRIOR FILING DATE: 1999-04-15
15 ; PRIOR APPLICATION NUMBER: 60/129,511
16 ; PRIOR FILING DATE: 1999-04-15
17 ; NUMBER OF SEQ ID NOS: 103
18 ; SOFTWARE: Microsoft Office 97
19 ; SEQ ID NO 3
20 ; LENGTH: 1378

```

TYPE: DNA  
ORGANISM: Dehalococcoides ethenogenes strain V/SFD  
US-10-061-071-3

Query Match 100.0%; Score 18; DB 9; Length 1378;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAGTGAAG 18  
112 GCTTCGGGAAGTGAAG 129

RESULT 7  
US-10-061-071-7  
Sequence 7, Application US/10061071  
Publication No. US20030077601A1  
GENERAL INFORMATION:  
APPLICANT: E.I. DUPONT DE NEMOURS & COMPANY  
APPLICANT: HENDRICKSON, EDWIN  
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS FOR THE IDENTIFICATION OF DECHLORINATING  
TITLE OF INVENTION: BACTERIA  
FILE REFERENCE: BC1002 US CIP  
CURRENT APPLICATION NUMBER: US/10/061,071  
CURRENT FILING DATE: 2002-01-29  
PRIOR APPLICATION NUMBER: US 60/129,511  
PRIOR FILING DATE: 1999-04-15  
PRIOR APPLICATION NUMBER: 60/129,511  
PRIOR FILING DATE: 1999-04-15  
NUMBER OF SEQ ID NOS: 103  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 7  
LENGTH: 1443  
TYPE: DNA  
ORGANISM: Dehalococcoides ethenogenes strain 195  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1353)..(1353)  
OTHER INFORMATION: N= unknown  
US-10-061-071-7

Query Match 100.0%; Score 18; DB 9; Length 1443;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAGTGAAG 18  
112 GCTTCGGGAAGTGAAG 129

RESULT 8  
US-09-878-574-6798  
Sequence 6798, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Thompson, Michael D.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants  
FILE REFERENCE: 38-21(15401)B  
CURRENT APPLICATION NUMBER: US/09/878,574  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 09/333,535  
PRIOR FILING DATE: 1999-06-14  
NUMBER OF SEQ ID NOS: 15775  
SEQ ID NO 6798  
LENGTH: 207  
TYPE: DNA  
ORGANISM: Glycine max  
OTHER INFORMATION: Clone ID: 701099017H1  
US-09-878-574-6798

Query Match 91.1%; Score 16.4; DB 10; Length 207;  
Best Local Similarity 94.4%; Pred. No. 17;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAGTGAAG 18  
40 GCTTCGGGAAGTGAAG 57

RESULT 9  
US-10-011-585A-14/C  
Sequence 14, Application US/10011585A  
Publication No. US2003039986A1  
GENERAL INFORMATION:  
APPLICANT: Sun, Yongming  
APPLICANT: Recipon, Hervé  
APPLICANT: Chen, Sei-Yu  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific  
TITLE OF INVENTION: Genes and Proteins  
FILE REFERENCE: DEX-0261  
CURRENT APPLICATION NUMBER: US/10/011,585A  
CURRENT FILING DATE: 2002-03-14  
PRIOR APPLICATION NUMBER: 60/245,740  
PRIOR FILING DATE: 2000-11-03  
NUMBER OF SEQ ID NOS: 245  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 14  
LENGTH: 244  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-011-585A-14

Query Match 85.6%; Score 15.4; DB 9; Length 244;  
Best Local Similarity 94.1%; Pred. No. 61;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAGTGAAG 17  
27 GCTTCGGGAAGTGAAG 11

RESULT 10  
US-09-918-995-36193  
Sequence 36193, Application US/09918995  
Publication No. US20030073623A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
FILE REFERENCE: 20411-756  
CURRENT APPLICATION NUMBER: US/09/918,995  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: US/09/235,076  
PRIOR FILING DATE: 1999-01-20  
NUMBER OF SEQ ID NOS: 38054  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 36193  
LENGTH: 459  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(459)  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-36193

Query Match 85.6%; Score 15.4; DB 9; Length 459;  
Best Local Similarity 94.1%; Pred. No. 64;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAGTGAAG 17  
112 GCTTCGGGAAGTGAAG 129

Db 112 GCTTCGGGAAACAGAG 128

## RESULT 11

US-10-102-524-314  
; Sequence 314, Application US/10102524  
; Publication No. US20030109434A1  
; GENERAL INFORMATION:  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Gordon, Brian  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER  
; FILE REFERENCE: 210121.572  
; CURRENT APPLICATION NUMBER: US/10/102,524  
; CURRENT FILING DATE: 2002-03-19  
; NUMBER OF SEQ ID NOS: 1863  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 314  
; LENGTH: 473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 318, 332, 427, 437, 443, 461, 463, 470  
; OTHER INFORMATION: n = A,T,C or G  
US-10-102-524-314

Query Match 85.6%; Score 15.4; DB 9; Length 473;  
Best Local Similarity 94.1%; Pred. No. 65;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAACAG 17  
|||||

Db 120 GCTTCGGGAAACAGAG 136

## RESULT 12

US-09-998-598-663  
; Sequence 663, Application US/09998598  
; Patent No. US20020150922A1  
; GENERAL INFORMATION:  
; APPLICANT: Stolk, John A.  
; APPLICANT: Xu, Jianshun  
; APPLICANT: Chenault, Ruth A.  
; APPLICANT: Mesgher, Madelein Joy  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.561  
; CURRENT APPLICATION NUMBER: US/09/998,598  
; CURRENT FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 2606  
; SOFTWARE: Corixa Invention Disclosure Database  
; SEQ ID NO 663  
; LENGTH: 507  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-998-598-663

Query Match 85.6%; Score 15.4; DB 10; Length 507;  
Best Local Similarity 94.1%; Pred. No. 65;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAACAG 17  
|||||

Db 120 GCTTCGGGAAACAGAG 136

## RESULT 13

US-10-011-585A-15  
; Sequence 15, Application US/10011585A

; Publication No. US20030039986A1

; GENERAL INFORMATION:  
; APPLICANT: Sun, Yongming  
; APPLICANT: Recipon, Hervé  
; APPLICANT: Chen, Sei-Yu  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific  
; TITLE OF INVENTION: Genes and Proteins  
; FILE REFERENCE: DEX-0261  
; CURRENT APPLICATION NUMBER: US/10/011,585A  
; CURRENT FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 60/245,740  
; PRIOR FILING DATE: 2000-11-03  
; NUMBER OF SEQ ID NOS: 245  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 531  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (202)  
; OTHER INFORMATION: a, c, g or t  
; NAME/KEY: unsure  
; LOCATION: (211)..(212)  
; OTHER INFORMATION: a, c, g or t  
; NAME/KEY: unsure  
; LOCATION: (217)  
; OTHER INFORMATION: a, c, g or t  
US-10-011-585A-15

Query Match 85.6%; Score 15.4; DB 9; Length 531;  
Best Local Similarity 94.1%; Pred. No. 65;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAACAG 17  
|||||

Db 505 GCTTCGGGAAACAGAG 521

## RESULT 14

US-09-864-761-9221  
; Sequence 9221, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: A60164-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263,6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665

```
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 9221
;; LENGTH: 560
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AP000041.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
US-09-864-761-9221

Query Match      85.6%; Score 15.4; DB 10; Length 560;
Best Local Similarity 94.1%; Pred. No. 65;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GCTTCGGGAACGTGAAG 17
        ||||| ||||| |||||
DB      64 GCTTCGGGAACGTGAAG 80

RESULT 15
US-09-864-761-26458
;; Sequence 26458, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aeomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
```

```
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 26458
;; LENGTH: 560
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AP000041.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
;; OTHER INFORMATION: NT HIT: AL161500.2, EVALUATE 1.40e+00
;; OTHER INFORMATION: EST_HUMAN HIT: BF446704.1, EVALUATE 2.10e+00
;; OTHER INFORMATION: SWISSPROT HIT: P38110, EVALUATE 6.30e+00
US-09-864-761-26458

Query Match      85.6%; Score 15.4; DB 10; Length 560;
Best Local Similarity 94.1%; Pred. No. 65;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GCTTCGGGAACGTGAAG 17
        ||||| ||||| |||||
DB      64 GCTTCGGGAACGTGAAG 80
```

Search completed: June 19, 2003, 08:44:28  
Job time : 158 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 04:28:14 ; Search time 51 Seconds

(without alignments)  
108.239 Million cell updates/sec

Title: US-09-911-860a-3

Sequence: 1 gcttcgggaactgaag 18

Scoring table: IDENTITY\_NUC

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, NA:\*

- 1: /cgn2\_6/ptodata/1/lna/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/lna/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/lna/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/lna/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/lna/CTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/lna/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.4	85.6	1647	3	US-09-101-146-44
2	15.4	85.6	6773	4	US-09-166-350-27
3	15.4	83.3	40352	3	US-08-846-111D-15
4	15.4	83.3	40352	4	US-09-443-077-15
5	14.8	82.2	4403765	4	US-09-103-840A-2
6	14.8	82.2	4411529	4	US-09-103-840A-1
7	14.4	80.0	87350	3	US-08-781-891-79
8	14.4	80.0	87543	4	US-09-791-211-3
9	14.4	80.0	4403765	4	US-09-103-840A-2
10	14.4	80.0	4411529	4	US-09-103-840A-1
11	13.8	76.7	219	1	US-08-700-575-21
12	13.8	76.7	219	1	US-07-807-043B-7
13	13.8	76.7	219	1	US-08-299-849B-7
14	13.8	76.7	219	2	US-08-142-368A-7
15	13.8	76.7	219	3	US-08-967-727-7
16	13.8	76.7	219	4	US-08-037-230D-7
17	13.8	76.7	219	1	US-08-465-167A-23
18	13.8	76.7	219	4	US-09-056-105-4
19	13.8	76.7	219	4	US-08-627-820-23
20	13.8	76.7	219	4	US-08-056-200-109
21	13.8	76.7	219	2	US-08-800-644-109
22	13.8	76.7	219	2	US-09-132-619-9
23	13.8	76.7	219	3	US-09-282-803B-9
24	13.8	76.7	219	4	US-09-510-654-9
25	13.8	76.7	219	4	US-09-061-709-8
26	13.8	76.7	219	4	US-09-061-709-6
27	13.8	76.7	219	1	US-08-530-492-3

c 28	13.8	76.7	3531	2	US-08-841-178-27	Sequence 27, Appl
c 29	13.8	76.7	3531	4	US-08-906-517-3	Sequence 3, Appl1
c 30	13.8	76.7	3531	2	US-08-841-178-26	Sequence 26, Appl
c 31	13.8	76.7	3881	1	US-08-299-953-2	Sequence 2, Appl1
c 32	13.8	76.7	3881	1	US-08-459-415-2	Sequence 2, Appl1
c 33	13.8	76.7	3881	1	US-09-066-687-2	Sequence 2, Appl1
c 34	13.8	76.7	3881	5	PCT-US95-11231-2	Sequence 2, Appl1
c 35	13.8	76.7	5674	1	US-07-807-043B-8	Sequence 8, Appl1
c 36	13.8	76.7	5674	1	US-08-190-411A-1	Sequence 1, Appl1
c 37	13.8	76.7	5674	2	US-08-299-849B-8	Sequence 8, Appl1
c 38	13.8	76.7	5674	1	US-08-560-024-1	Sequence 1, Appl1
c 39	13.8	76.7	5674	2	US-08-142-368A-8	Sequence 8, Appl1
c 40	13.8	76.7	5674	3	US-08-967-727-8	Sequence 8, Appl1
c 41	13.8	76.7	5674	4	US-08-037-230D-8	Sequence 8, Appl1
c 42	13.8	76.7	6152	1	US-08-557-139-1	Sequence 8, Appl1
c 43	13.8	76.7	6803	3	US-08-665-259-19	Sequence 19, Appl
c 44	13.8	76.7	6803	3	US-08-762-500-19	Sequence 19, Appl
c 45	13.8	76.7	65042	4	US-09-784-316-3	Sequence 3, Appl1

## ALIGNMENTS

RESULT 1  
US-09-101-146-44

Sequence 44, Application US/09101146

Patent No. 6124125

GENERAL INFORMATION:

APPLICANT: Dartmouth College, St. Vincents Institute of

APPLICANT: Medical Research, Kemp et al.

TITLE OF INVENTION: NO. 6124125el AMP Activated Protein Kinase

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESSES:

ADDRESSER: Jane Massey Licata, Esq.

STREET: 66 E. Main Street

City: Marlton

STATE: NJ

COUNTRY: USA

ZIP: 08053

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE

COMPUTER: IBM PC

OPERATING SYSTEM: WINDOWS 95

SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/101,146

FILING DATE: October 7, 1998

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PN7450

FILING DATE: 8 JAN 1996

ATTORNEY/AGENT INFORMATION:

NAME: Jane Massey Licata

REGISTRATION NUMBER: 32,257

REFERENCE/DOCKET NUMBER: DC-0050

TELECOMMUNICATION INFORMATION:

TELEPHONE: (856) 810-1515

TELEFAX: (856) 810-1454

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 1647

TYPE: Nucleic acid

STRANDEDNESS: Single

TOPOLOGY: Linear

ANTI-SENSE: NO

US-09-101-146-44

Query Match

Best Local Similarity 94.1%

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Score 15.4; DB 3; Length 1647;

Pred. No. 32;

Score 15.4; DB 3; Length 1647;

Pred. No. 32;

Score 15.4; DB 3; Length 1647;

Pred. No. 32;

Score 15.4; DB 3; Length 1647;

Pred. No. 32;

Score 15.4; DB 3; Length 1647;

Pred. No. 32;

Score 15.4; DB 3; Length 1647;

Pred. No. 32;

Score 15.4; DB 3; Length 1647;

Db 78 CTTGGGAAAGTGAGG 94

## RESULT 2

US-09-166-350-27  
Sequence 27, Application US/09166350A  
Patent No. 6440663  
GENERAL INFORMATION:  
APPLICANT: Scanlan, Matthew  
APPLICANT: Chen, Yao  
APPLICANT: Stockert, Elisabeth  
APPLICANT: Old, Lloyd  
APPLICANT: Jager, Elke  
APPLICANT: Knuth, Alex  
TITLE OF INVENTION: Renal Cancer Associated Antigens and  
TITLE OF INVENTION: Uses Therefor  
FILE REFERENCE: L0461/7051  
CURRENT APPLICATION NUMBER: US/09/166,350A  
CURRENT FILING DATE: 1998-10-05  
EARLIER APPLICATION NUMBER: US 09/166,350  
EARLIER FILING DATE: 1998-10-05  
NUMBER OF SEQ. ID NOS.: 35  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 27  
LENGTH: 6773  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-166-350-27

Query Match 85.6%; Score 15.4; DB 4; Length 6773;  
Best Local Similarity 94.1%; Pred. No. 38;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCTTGGGAAAGTGAG 17

Db 4684 GCTTGGGAAAGTGAG 4700

## RESULT 3

US-08-846-111D-15  
Sequence 15, Application US/08846111D  
Patent No. 6017705  
GENERAL INFORMATION:  
APPLICANT: Lurguin, Christophe; Brasseur, Francis;  
APPLICANT: Boon-Falleur, Thierry  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Are  
TITLE OF INVENTION: Members Of The Mage-B Family and Uses Thereof  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022-7513  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect For DOS 6.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/846,111D  
FILING DATE: 25-APRIL-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/658,578  
FILING DATE: 5-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/403,388  
FILING DATE: 14-MARCH-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6017705man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5444.1

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ. ID NO.: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40352 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-846-111D-15

Query Match 83.3%; Score 15; DB 3; Length 40352;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CTTGGGAAAGTGAA 16

Db 32274 CTTGGGAAAGTGAA 32288

## RESULT 4

US-09-443-077-15  
Sequence 15, Application US/09443077  
Patent No. 6392016  
GENERAL INFORMATION:  
APPLICANT: Lurguin, Christophe; Brasseur, Francis;  
APPLICANT: Boon-Falleur, Thierry  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Are  
TITLE OF INVENTION: Members Of The Mage-B Family and Uses Thereof  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022-7513  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect For DOS 6.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/443,077  
FILING DATE: 26-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/846,111  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/403,388  
FILING DATE: 14-MARCH-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6392016man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5444.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ. ID NO.: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40352 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-09-443-077-15

Query Match 83.3%; Score 15; DB 4; Length 40352;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CTTGGGAAAGTGAA 16

Db 32274 CTTGGGAAAGTGAA 32288

```

RESULT 5
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2
Query Match      82.2%; Score 14.8; DB 4; Length 4403765;
Best Local Similarity 88.9%; Pred. NO. 64;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

OY      1 GCTTCGGGAACCTGAAGC 18
          || ||||| |||||
Db      1603115 GCATCGGACACTGAAGC 1603098

RESULT 6
US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1
Query Match      82.2%; Score 14.8; DB 4; Length 4411529;
Best Local Similarity 88.9%; Pred. NO. 64;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

OY      1 GCTTCGGGAACCTGAAGC 18
          || ||||| |||||
Db      1603280 GCATCGGACACTGAAGC 1603263

RESULT 7
US-08-781-891-79/c
Sequence 79, Application US/08781891
Patent No. 6090620

```

```

GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 87350 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-79

Query Match 80.0%; Score 14.4; DB 3: Length 87350;
Best Local Similarity 93.8%; Pred. No. 1,7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 CTTGGGAAGTGAAG 17
      ||| |||||
Db 76424 CTTGGGAAGTGAAG 76409

RESULT 8
US-09-791-211-3/C
Sequence 3, Application US/09791211
Patent No. 6448080
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
APPLICANT: Andrew T. Malt
TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
FILE REFERENCE: RTS-0205
CURRENT APPLICATION NUMBER: US/09/791,211
CURRENT FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 90
SEQ ID NO 3
LENGTH: 87543
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 7421
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 7427
OTHER INFORMATION: unknown
NAME/KEY: unsure

```



Page 4

```

: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 52786
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 52787
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 53384
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 54684
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 59215
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 59235
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 59242
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 63290
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 66614
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 68660
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 68697
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 68718
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 68733
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 68739
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 69785
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 79134
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 79198
: OTHER INFORMATION: unknown
: NAME/KEY: unsure
: LOCATION: 86336
: OTHER INFORMATION: unknown
: OTHER INFORMATION: unknown
:
: US-09-791-211-3
:
: Query Match      80.0%; Score 14.4; DB 4; Length 87543;
: Best Local Similarity 93.8%; Pred. No. 1.7e+02;
: Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
:
OY      2 CTTGGGAACTGAAG 17
      ||| |||||
Db      76617 CTTGGGAACTGAAG 76602

```

```

; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match      80.0%; Score 14.4; DB 4; Length 4403765;
Best Local Similarity 93.8%; Pred. No. 68;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      3 TTCCGGAAACTGAAG 18
DB      429374 TTCCGGAAACTGAATG 429389

RESULT 10
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match      80.0%; Score 14.4; DB 4; Length 4411529;
Best Local Similarity 93.8%; Pred. No. 68;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      3 TTCCGGAAACTGAAG 18
DB      429291 TTCCGGAAACTGAATG 429306
```

```

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,575
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J
; REGISTRATION NUMBER: 36749
; REFERENCE/DOCKET NUMBER: SP-100 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 219 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: Rheumatoid Synovium
; CLONE: 81865
US-08-700-575-21

Query Match      76.7%; Score 13.8; DB 1; Length 219;
Best Local Similarity 88.2%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      2 CTTCGGAAACTGAAG 18
DB      75 CTTCGGAAACTGAAG 91

RESULT 12
US-07-807-043B-7
; Sequence 7, Application US/07807043B
; Patent No. 5342774
; GENERAL INFORMATION:
; APPLICANT: BOON, Thierry, Van den Eynde, Beno t
; TITLE OF INVENTION: Tumor Rejection Antigen Precursors, Tumor
; TITLE OF INVENTION: Rejection Antigens and Uses Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/807,043B
; FILING DATE: 19911212
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/764,364
; FILING DATE: 23-SEPTEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/728,838
; FILING DATE: 9-JULY-1991
```

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/705,702  
FILING DATE: 23-May-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5342774man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 253.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2419 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: singular  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
US-07-807-0438-7

Query Match 76.7%; Score 13.8; DB 1; Length 2419;  
Best Local Similarity 88.2%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACTGTAAG 17  
|||||  
Db 2361 GCTTGGGAACTGCAG 2377

RESULT 13  
US-08-299-849B-7  
Sequence 7, Application US/08299849B  
Patent No. 5612201  
GENERAL INFORMATION:  
APPLICANT: De Plaen, Etienne; Boon-Falleur, Thierry;  
APPLICANT: Leth, Bernard; Szikora, Jean-Pierre; De Smet, Charles;  
APPLICANT: Chomez, Patrick  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful In  
TITLE OF INVENTION: Determining Expression Of A Tumor Antigen Precursor  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Felle & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/299, 849B  
FILING DATE: 1-SEPTEMBER-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/037, 230  
FILING DATE: 26-MARCH-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/04354  
FILING DATE: 22-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/807, 043  
FILING DATE: 12-DECEMBER-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/764, 364  
FILING DATE: 23-SEPTEMBER-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/728, 838  
FILING DATE: 9-JULY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/705, 702  
FILING DATE: 23-May-1991  
ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5612201man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5355  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2419 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
US-08-299-849B-7

Query Match 76.7%; Score 13.8; DB 1; Length 2419;  
Best Local Similarity 88.2%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACTGTAAG 17  
|||||  
Db 2361 GCTTGGGAACTGCAG 2377

RESULT 14  
US-08-142-368A-7  
Sequence 7, Application US/08142368A  
Patent No. 5925729  
GENERAL INFORMATION:  
APPLICANT: Boon-Falleur, Thierry; Van der Bruggen, Thierry;  
APPLICANT: Van den Eynde, Beno t; Van Pel, Aline; De Plaen, Etienne;  
APPLICANT: Lurquin, Christophe; Chomez, Patrick; Traversari, Catia  
TITLE OF INVENTION: Tumor Rejection Antigen Precursors, Tumor  
TITLE OF INVENTION: Rejection Antigens and Uses Thereof  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Felle & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/142, 368A  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/04354  
FILING DATE: 22-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/807, 043  
FILING DATE: 12-DECEMBER-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/764, 364  
FILING DATE: 23-SEPTEMBER-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/728, 838  
FILING DATE: 9-JULY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/705, 702  
FILING DATE: 23-May-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5925729man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5253.4-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 7:



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 07:34:16 ; Search time 1864 Seconds

(without alignments)  
281.036 Million cell updates/sec

Title: US-09-911-860A-3

Sequence: 1 gcttcggaactgaag 18

Scoring table: IDENTITY-NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_iny:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	100.0	18	6	AX356919	AX356919 Sequence
2	18	100.0	501	1	AF093598	AF093598 Freshwater
3	18	100.0	584	1	AF427935	AF427935 Unculture
4	18	100.0	685	1	AF447150	AF447150 Unculture
5	18	100.0	777	1	AF388544	AF388544 Unculture
6	18	100.0	1212	6	AX039535	AX039535 Sequence
7	18	100.0	1212	6	AX039537	AX039537 Sequence
8	18	100.0	1212	6	AX039538	AX039538 Sequence
9	18	100.0	1212	6	AX039539	AX039539 Sequence
10	18	100.0	1286	1	AF427908	AF427908 Unculture
11	18	100.0	1315	1	AF388543	AF388543 Unculture
12	18	100.0	1332	1	UBA249262	AJ249262 Bacterium
13	18	100.0	1331	1	AF427937	AF427937 Unculture
14	18	100.0	1335	6	AX039536	AX039536 Sequence
15	18	100.0	1377	1	AF388531	AF388531 Unculture
16	18	100.0	1377	1	AF388532	AF388532 Unculture
17	18	100.0	1377	1	AF388533	AF388533 Unculture
18	18	100.0	1377	1	AF388534	AF388534 Unculture
19	18	100.0	1377	1	AF388535	AF388535 Unculture
20	18	100.0	1377	1	AF388536	AF388536 Unculture
21	18	100.0	1377	1	AF388537	AF388537 Unculture
22	18	100.0	1377	1	AF388540	AF388540 Unculture
23	18	100.0	1377	1	AF388541	AF388541 Unculture
24	18	100.0	1377	1	AF388542	AF388542 Unculture
25	18	100.0	1377	1	AF388546	AF388546 Unculture
26	18	100.0	1377	1	AF388548	AF388548 Unculture
27	18	100.0	1377	1	AF388549	AF388549 Unculture
28	18	100.0	1378	1	AF388530	AF388530 Unculture
29	18	100.0	1378	1	AF388539	AF388539 Unculture
30	18	100.0	1378	1	AF388550	AF388550 Unculture
31	18	100.0	1398	1	AF427910	AF427910 Unculture
32	18	100.0	1409	1	AF529119	AF529119 Unculture
33	18	100.0	1421	1	AF230641	AF230641 Bacterium
34	18	100.0	1423	1	AF427907	AF427907 Unculture
35	18	100.0	1426	1	AF427912	AF427912 Unculture
36	18	100.0	1426	1	AF529124	AF529124 Unculture
37	18	100.0	1434	1	AF004928	AF004928 Dehalococ
38	18	100.0	1435	1	AF357918	AF357918 Dehalococ
39	18	100.0	1443	6	AX039540	AX039540 Sequence
40	18	100.0	1455	1	AF427909	AF427909 Unculture
41	18	100.0	1456	1	P231182	AJ231182 Planctomy
42	18	100.0	1522	1	AB015586	AB015586 Unidentif
43	18	100.0	1532	8	AC010704	AC010704 Arabidops
44	17	94.4	384	1	UNSSBR2093	X84623 Unknown org
45	17	94.4	912	1	ARAJ3138	AJ003138 unidentif

#### ALIGNMENTS

RESULT 1  
AX356919  
LOCUS AX356919 18 bp DNA linear PAT 13-FEB-2002  
DEFINITION Sequence 3 from Patent EP116216.  
ACCESSION AX356919  
VERSION AX356919.1 GI:18674118  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
Nakamura,K.C. and Ueno,T.C.  
Nucleic acid, nucleic acid for detecting chlorinated  
ethylene-decomposing bacteria, probe, method of detecting  
chlorinated ethylene-decomposing bacteria, and method of

decomposing chlorinated ethylene or ethane  
Patent: EP 1176216-A 3 30-JAN-2002;  
Kurita Water Industries Ltd. (JP)  
Location/Qualifiers

JOURNAL  
FEATURES  
SOURCE

BASE COUNT 5 a 3 c 7 g 3 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 18;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACGTGAAG 18  
DB 1 GCTTCGGGAACGTGAAG 18

RESULT 2  
AF093598  
LOCUS AF093598 501 bp RNA linear BCT 27-AUG-1999  
DEFINITION Freshwater sediment clone ppd17rna 16S ribosomal RNA, partial  
ACCESSION AF093598  
VERSION AF093598  
KEYWORDS GI:4585198  
SOURCE  
ORGANISM freshwater sediment clone ppd17rna.  
Bacteria; environmental samples.  
REFERENCE 1 (bases 1 to 501)  
AUTHORS Miskin, I.P., Farrimond, P. and Head, I.M.  
TITLE Identification of novel bacterial lineages as active members of  
microbial populations in a freshwater sediment using a rapid RNA  
extraction procedure and RT-PCR  
JOURNAL Microbiology 145 (Pt 8), 1977-1987 (1999)  
MEDLINE 99392460  
PUBMED 10463164

REFERENCE 2 (bases 1 to 501)  
AUTHORS Miskin, I.P. and Head, I.M.  
TITLE Rapid extraction of indigenous bacterial RNAs from freshwater  
sediments and reverse transcriptase-polymerase chain reaction  
amplification of ribosomal and messenger RNA  
unpublished

JOURNAL 3 (bases 1 to 501)  
REFERENCE Miskin, I.P. and Head, I.M.  
AUTHORS Direct Submission  
TITLE Submitted (23-SEP-1998) Fossil Fuels and Environmental Geochemistry  
JOURNAL (Postgraduate Institute), University of Newcastle, Devonshire  
Terrace, Newcastle-upon-Tyne NE1 7RU, UK  
Location/Qualifiers

FEATURES  
SOURCE 1. 501  
/organism="freshwater sediment clone ppd17rna"  
/db\_xref="taxon:91597"  
/clone="ppd17rna"  
/note="Isolated from ribosomal RNA template from the  
sediments of the eutrophic freshwater lake Priest Pot;  
sample depth 11-12 cm"

RRNA  
BASE COUNT 119 a 120 c 169 g 92 t 1 others  
ORIGIN

Query Match 100.0%; Score 18; DB 1; Length 501;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACGTGAAG 18  
DB 145 GCTTCGGGAACGTGAAG 162

RESULT 3  
AF427935 584 bp DNA linear BCT 16-JUL-2002  
LOCUS AF427935  
DEFINITION Uncultured bacterium clone 2B 16S ribosomal RNA gene, partial  
sequence.  
AF427935  
AF427935.1 GI:16566620

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM uncultured bacterium.  
Bacteria; environmental samples.

REFERENCE 1 (bases 1 to 584)  
AUTHORS Richardson, R.E., Bhupathiraju, V.K., Song, D.L., Goulet, T.A. and  
Alvarez-Cohen, L.  
TITLE Phylogenetic characterization of microbial communities that  
reductively dechlorinate TCE based upon a combination of molecular  
techniques  
Evitron. Sci. Technol. 36 (12), 2652-2662 (2002)

JOURNAL 22093797  
MEDLINE 12099461  
PUBMED 2 (bases 1 to 584)  
REFERENCE Richardson, R.E., Bhupathiraju, V.K., Song, D.L., Goulet, T.A. and  
AUTHORS Alvarez-Cohen, L.  
TITLE Direct Submission  
JOURNAL Submitted (04-OCT-2001) Civil and Environmental Engineering,  
University of California at Berkeley, 631 Davis Hall, Berkeley, CA  
94720-1710, USA  
Location/Qualifiers

FEATURES  
SOURCE 1. 584  
/organism="uncultured bacterium"  
/db\_xref="taxon:77133"  
/clone="2B"  
/note="RFLP type 2; from TCE-dechlorinating consortium  
ANAS II"

RNA  
BASE COUNT 147 a 119 c 181 g 136 t 1 others

Query Match 100.0%; Score 18; DB 1; Length 584;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACGTGAAG 18  
DB 95 GCTTCGGGAACGTGAAG 112

RESULT 4  
AF447150 685 bp DNA linear BCT 12-DEC-2001  
LOCUS AF447150  
DEFINITION Uncultured bacterium clone TDC-S1:33 16S ribosomal RNA gene,  
partial sequence.  
AF447150  
AF447150.1 GI:17530460

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM uncultured bacterium.  
Bacteria; environmental samples.

REFERENCE 1 (bases 1 to 685)  
AUTHORS Dennis, P.C., Sleep, B.E., Fulthorpe, R.R. and Liss, S.N.  
TITLE Phylogenetic Characterization of an Anaerobic Bacterial Consortium  
Capable of Degradation Saturation Concentrations of Tetrachloroethene  
unpublished  
JOURNAL 2 (bases 1 to 685)  
REFERENCE Dennis, P.C., Sleep, B.E., Fulthorpe, R.R. and Liss, S.N.  
AUTHORS Direct Submission  
TITLE Submitted (11-NOV-2001) Department of Chemistry, Biology and  
JOURNAL Chemical Engineering, Ryerson Polytechnic University, 350 Victoria  
Street, Toronto, Ontario M5B 2K3, Canada  
Location/Qualifiers

FEATURES  
SOURCE 1. 685  
/organism="uncultured bacterium"

/db\_xref="taxon:77133"  
/clone="TDC-S1.33"  
<1. >685  
/product="16S ribosomal RNA"

BASE COUNT 182 a 134 c 219 g 149 t 1 others

ORIGIN

Query Match 100.0%; Score 18; DB 1; Length 685;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAACTGAAG 18  
|||||

Db 63 GCTTCGGGAAACTGAAG 80

RESULT 5  
AF388544 777 bp DNA linear BCT 29-JAN-2002  
LOCUS Uncultured Dehalococcoides sp. clone DHC-nfdpc 16S ribosomal RNA  
DEFINITION  
ACCESSION AF388544  
VERSION AF388544.1 GI:18389817  
KEYWORDS uncultured Dehalococcoides sp.  
SOURCE uncultured Dehalococcoides sp.  
ORGANISM Bacteria; Dehalococcoides group; Dehalococcoides; environmental samples.

REFERENCE 1 (bases 1 to 777)  
Hendrickson, E.R., Payne, J.A., Young, R.M., Starr, M.G., Perry, M.P., Fhnestock, S., Ellis, D.E. and Ebersole, R.C.  
Molecular Analysis of Dehalococcoides 16S Ribosomal DNA from Chloroethene-Contaminated Sites throughout North America and Europe Appl. Environ. Microbiol. 68 (2), 485-495 (2002)

JOURNAL 11823182  
TITLE 2 (bases 1 to 777)  
AUTHORS Hendrickson, E.R., Payne, J.A., Young, R.M., Starr, M.G., Perry, M.P., Fhnestock, S., Ellis, D.E. and Ebersole, R.C.  
REFERENCE Direct Submission  
JOURNAL Submitted (06-JUN-2001) Central R & D/ Corporate Center for Engineering Research, E.I. Du Pont de Nemours and Company, P.O. Box 6101, Glasgow Site, 300/318, Newark, DE 19714-6101, USA  
TITLE Location/Qualifiers

FEATURES  
source 1..777  
/organism="uncultured Dehalococcoides sp."  
/db\_xref="taxon:171952"  
/clone="DHC-nfdpc"  
/country="USA: Niagara Falls, NY"  
/note="Isolated from chloroethene-contaminated site"  
<1. >777  
/product="16S ribosomal RNA"

BASE COUNT 208 a 154 c 244 g 171 t

ORIGIN

Query Match 100.0%; Score 18; DB 1; Length 777;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAACTGAAG 18  
|||||

Db 112 GCTTCGGGAAACTGAAG 129

RESULT 6  
AX039535 1212 bp DNA linear PAT 18-NOV-2000  
LOCUS Sequence 2 from Patent W00063443.  
DEFINITION  
ACCESSION AX039535  
VERSION AX039535.1 GI:11229557  
KEYWORDS Dehalococcoides ethenogenes.  
SOURCE Dehalococcoides ethenogenes  
ORGANISM Bacteria; Dehalococcoides group; Dehalococcoides.

REFERENCE 1 (bases 1 to 1212)  
AUTHORS Hendrickson, E.R. and Ebersole, R.C.  
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria  
JOURNAL Patent: WO 0063443-A 2 26-OCT-2000;  
E.I. DU PONT DE NEMOURS AND COMPANY (US)  
FEATURES  
source 1..1212  
/organism="Dehalococcoides ethenogenes"  
/db\_xref="taxon:61435"

BASE COUNT 321 a 252 c 369 g 269 t 1 others

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1212;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAACTGAAG 18  
|||||

Db 112 GCTTCGGGAAACTGAAG 129

RESULT 7  
AX039537 1212 bp DNA linear PAT 18-NOV-2000  
LOCUS Sequence 4 from Patent W00063443.  
DEFINITION  
ACCESSION AX039537  
VERSION AX039537.1 GI:11229559  
KEYWORDS Dehalococcoides ethenogenes.  
SOURCE Dehalococcoides ethenogenes.  
ORGANISM Bacteria; Dehalococcoides group; Dehalococcoides.

REFERENCE 1 (bases 1 to 1212)  
Hendrickson, E.R. and Ebersole, R.C.  
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria  
JOURNAL Patent: WO 0063443-A 4 26-OCT-2000;  
E.I. DU PONT DE NEMOURS AND COMPANY (US)  
FEATURES  
source 1..1212  
/organism="Dehalococcoides ethenogenes"  
/db\_xref="taxon:61435"

BASE COUNT 318 a 253 c 372 g 269 t

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1212;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAAACTGAAG 18  
|||||

Db 112 GCTTCGGGAAACTGAAG 129

RESULT 8  
AX039538 1212 bp DNA linear PAT 18-NOV-2000  
LOCUS Sequence 5 from Patent W00063443.  
DEFINITION  
ACCESSION AX039538  
VERSION AX039538.1 GI:11229560  
KEYWORDS Dehalococcoides ethenogenes.  
SOURCE Dehalococcoides ethenogenes.  
ORGANISM Bacteria; Dehalococcoides group; Dehalococcoides.

REFERENCE 1 (bases 1 to 1212)  
Hendrickson, E.R. and Ebersole, R.C.  
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria  
JOURNAL Patent: WO 0063443-A 5 26-OCT-2000;  
E.I. DU PONT DE NEMOURS AND COMPANY (US)  
FEATURES  
source 1..1212  
/organism="Dehalococcoides ethenogenes"

BASE COUNT 320 a 253 c 371 g 268 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1212;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTCGGGAACTGAAGC 18  
|||||  
DB 112 GCTTCGGGAACTGAAGC 129

RESULT 9  
AX039539 1212 bp DNA Linear PAT 18-NOV-2000  
LOCUS Sequence 6 from Patent WO0063443.  
ACCESSION AX039539  
VERSION AX039539.1 GI:11229561  
KEYWORDS

SOURCE Dehalococcoides ethenogenes.  
Dehalococcoides ethenogenes

REFERENCE 1 (bases 1 to 1212)  
Hendrickson, E.R. and Ebersole, R.C.  
Nucleic acid fragments for the identification of dechlorinating ba

TITLE Patent: WO 0063443-A 6 26-OCT-2000;  
PATENT: DU PONT DE NEMOURS AND COMPANY (US)  
LOCATION/Qualifiers

FEATURES  
SOURCE 1. 1212  
/organism="Dehalococcoides ethenogenes"  
/db\_xref="taxon:61435"

BASE COUNT 320 a 253 c 371 g 268 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1212;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTCGGGAACTGAAGC 18  
|||||  
DB 112 GCTTCGGGAACTGAAGC 129

RESULT 10  
AF427908 1286 bp DNA Linear BCT 16-JUL-2002  
LOCUS Uncultured bacterium clone 56 16S ribosomal RNA gene, partial  
DEFINITION  
SEQUENCE  
ACCESSION AF427908  
VERSION AF427908.1 GI:16566579  
KEYWORDS

SOURCE uncultured bacterium.  
uncultured bacterium

REFERENCE 1 (bases 1 to 1286)  
Richardson, R.E., Bhupathiraju, V.K., Song, D.L., Goulet, T.A. and  
Alvarez-Cohen, L.

TITLE Phylogenetic characterization of microbial communities that  
reductively dechlorinate TCE based upon a combination of molecular  
techniques

JOURNAL Environ. Sci. Technol. 36 (12), 2652-2662 (2002)  
MEDLINE 22093797  
PUBMED 12099461

REFERENCE 2 (bases 1 to 1286)  
Richardson, R.E., Bhupathiraju, V.K., Song, D.L., Goulet, T.A. and  
Alvarez-Cohen, L.

TITLE Direct Submission  
Submitted (04-OCT-2001) Civil and Environmental Engineering,  
University of California at Berkeley, 631 Davis Hall, Berkeley, CA  
94720-1710, USA

Location/Qualifiers

KEYWORDS  
SOURCE  
ORIGIN

source 1. 1286  
/organism="uncultured bacterium"  
/db\_xref="taxon:77133"

/clone="56"  
/note="RFLP type 4; from TCE-dechlorinating consortium  
ANAS I"

<1. >1286  
/product="16S ribosomal RNA"

BASE COUNT 338 a 274 c 385 g 288 t 1 others  
ORIGIN

Query Match 100.0%; Score 18; DB 1; Length 1286;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTCGGGAACTGAAGC 18  
|||||  
DB 61 GCTTCGGGAACTGAAGC 78

RESULT 11  
AF388543 1315 bp DNA Linear BCT 29-JAN-2002  
LOCUS Uncultured Dehalococcoides sp. clone DHC-1n2 16S ribosomal RNA  
DEFINITION  
ACCESSION AF388543  
VERSION AF388543.1 GI:18389816  
KEYWORDS

SOURCE uncultured Dehalococcoides sp.  
uncultured Dehalococcoides sp.

ORGANISM Bacteria; Dehalococcoides group; Dehalococcoides; environmental  
samples.

REFERENCE 1 (bases 1 to 1315)  
Hendrickson, E.R., Payne, J.A., Young, R.M., Starr, M.G., Perry, M.P.,  
Fahnestock, S., Ellis, D.E. and Ebersole, R.C.  
Molecular Analysis of Dehalococcoides 16S Ribosomal DNA from  
Chloroethene-Contaminated Sites throughout North America and Europe  
Appl. Environ. Microbiol. 68 (2), 485-495 (2002)  
11823182

REFERENCE 2 (bases 1 to 1315)  
Hendrickson, E.R., Payne, J.A., Young, R.M., Starr, M.G., Perry, M.P.,  
Fahnestock, S., Ellis, D.E. and Ebersole, R.C.  
Direct Submission  
Submitted (06-JUN-2001) Central R & D/Corporate Center for  
Engineering Research, E.I. Du Pont de Nemours and Company, P.O. Box  
6101, Glasgow Site, 300/318, Newark, DE 19714-6101, USA

LOCATION/Qualifiers

FEATURES  
SOURCE 1. 1315  
/organism="uncultured Dehalococcoides sp."  
/db\_xref="taxon:171952"

/clone="DHC-1n2"  
/country="USA: Lorenz, CA"  
/note="Isolated from chloroethene-contaminated site"

<1. >1315  
/product="16S ribosomal RNA"

BASE COUNT 343 a 279 c 401 g 292 t  
ORIGIN

Query Match 100.0%; Score 18; DB 1; Length 1315;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTCGGGAACTGAAGC 18  
|||||  
DB 112 GCTTCGGGAACTGAAGC 129

RESULT 12  
UBA249262 1322 bp DNA Linear BCT 28-AUG-2000  
LOCUS Bacterium DCEH2 16S rRNA gene.  
DEFINITION  
ACCESSION UBA249262  
VERSION UBA249262.1 GI:9944249

Location/Qualifiers



KEYWORDS 16S ribosomal RNA: 16S rRNA gene.  
SOURCE bacterium DCEH2.  
ORGANISM Bacterium DCEH2

REFERENCE 1 (bases 1 to 1322)  
AUTHORS Windfuhr,C., Mau,M., Scholz-Muramatsu,H. and Diekert,G.  
TITLE Anaerobic reductive dechlorination of chlorinated ethenes with an enriched mixed culture - Physiological characterization and community analysis  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1322)  
AUTHORS Windfuhr,C.  
TITLE Direct Submission  
JOURNAL Submitted (31-AUG-1999) Windfuhr C., Institute for Sanitary Engineering, Department of Biology, University of Stuttgart, Bandtlae 2, 70569 Stuttgart, 70569, GERMANY

FEATURES  
source Location/Qualifiers  
1. .1322  
/organism="bacterium DCEH2"  
/db\_xref="taxon:103252"  
/clone="DCEH2"  
/country="Germany"  
/note="clone from dechlorinating enrichment mixed culture"  
gene 1. .1322  
/gene="16S rRNA"  
rRNA 1. .1322  
/gene="16S rRNA"  
/product="16S ribosomal RNA"  
BASE COUNT 346 a 280 c 399 g 294 t 3 others  
ORIGIN

Query Match 100.0%; Score 18; DB 1; Length 1322;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACGTGAGG 18  
DB 126 GCTTCGGGAACGTGAGG 143

RESULT 13  
AF427937  
LOCUS AF427937 1331 bp DNA linear BCT 16-JUL-2002  
DEFINITION Uncultured bacterium clone 5H 16S ribosomal RNA gene, partial sequence.  
ACCESSION AF427937  
VERSION AF427937.1 GI:16566623  
KEYWORDS uncultured bacterium.  
SOURCE uncultured bacterium.  
ORGANISM Bacteria; environmental samples.  
REFERENCE 1 (bases 1 to 1331)  
AUTHORS Richardson,R.E., Bhupathiraju,V.K., Song,D.L., Goulet,T.A. and Alvarez-Cohen,L.  
TITLE Phylogenetic characterization of microbial communities that reductively dechlorinate TCE based upon a combination of molecular techniques  
JOURNAL Environ. Sci. Technol. 36 (12), 2652-2662 (2002)  
MEDLINE 22093797  
PUBMED 12099461  
REFERENCE 2 (bases 1 to 1331)  
AUTHORS Richardson,R.E., Bhupathiraju,V.K., Song,D.L., Goulet,T.A. and Alvarez-Cohen,L.  
TITLE Direct Submission  
JOURNAL Submitted (04-OCT-2001) Civil and Environmental Engineering, University of California at Berkeley, 631 Davis Hall, Berkeley, CA 94720-1710, USA

FEATURES  
source Location/Qualifiers  
1. .1331  
/organism="uncultured bacterium"  
/db\_xref="taxon:77133"  
/clone="5H"  
/note="RFLP type 22; from TCE-dechlorinating consortium"

ANAS 11"  
rRNA <1..>1331  
/product="16S ribosomal RNA"  
BASE COUNT 345 a 291 c 403 g 292 t  
ORIGIN

Query Match 100.0%; Score 18; DB 1; Length 1331;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACGTGAGG 18  
DB 37 GCTTCGGGAACGTGAGG 54

RESULT 14  
AX039536  
LOCUS AX039536 1335 bp DNA linear PAT 18-NOV-2000  
DEFINITION Sequence 3 from Patent W00063443.  
ACCESSION AX039536  
VERSION AX039536.1 GI:11229558  
KEYWORDS Dehalococcoides ethenogenes.  
SOURCE Dehalococcoides ethenogenes.  
ORGANISM Dehalococcoides ethenogenes.  
REFERENCE 1 (bases 1 to 1335)  
AUTHORS Hendrickson,E.R. and Ebersole,R.C.  
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria  
JOURNAL Patent: WO 0063443-A 3 26-OCT-2000;  
ORIGIN E.I. DU PONT DE NEMOURS AND COMPANY (US)

FEATURES  
source Location/Qualifiers  
1. .1335  
/organism="Dehalococcoides ethenogenes"  
/db\_xref="taxon:61435"  
BASE COUNT 348 a 288 c 403 g 296 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1335;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACGTGAGG 18  
DB 112 GCTTCGGGAACGTGAGG 129

RESULT 15  
AF388531  
LOCUS AF388531 1377 bp DNA linear BCT 29-JAN-2002  
DEFINITION Uncultured Dehalococcoides sp. clone DHC-bmtc 16S ribosomal RNA gene, partial sequence.  
ACCESSION AF388531  
VERSION AF388531.1 GI:18389804  
KEYWORDS uncultured Dehalococcoides sp.  
SOURCE uncultured Dehalococcoides sp.  
ORGANISM Bacteria; Dehalococcoides group; Dehalococcoides; environmental samples.  
REFERENCE 1 (bases 1 to 1377)  
AUTHORS Hendrickson,E.R., Payne,J.A., Young,R.M., Starr,M.G., Perry,M.P., Fahnestock,S., Ellis,D.E. and Ebersole,R.C.  
TITLE Molecular Analysis of Dehalococcoides 16S Ribosomal DNA from Chloroethene-Contaminated Sites throughout North America and Europe Appl. Environ. Microbiol. 68 (2), 485-495 (2002)  
PUBMED 11823182  
REFERENCE 2 (bases 1 to 1377)  
AUTHORS Hendrickson,E.R., Payne,J.A., Young,R.M., Starr,M.G., Perry,M.P., Fahnestock,S., Ellis,D.E. and Ebersole,R.C.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUN-2001) Central R & D/Corporate Center for Engineering Research, E.I. Du Pont de Nemours and Company, P.O. Box 6101, Glasgow Site, 300/318, Newark, DE 19714-6101, USA

FEATURES  
source  
Location/Qualifiers  
1. 1377  
/organism="uncultured Dehalococcoides sp."  
/db\_xref="taxon:171952"  
/clone="DHC-Dmtc"  
/country="USA: Beaumont, TX"  
/note="Isolated from chloroethene-contaminated sludge pond"  
<1. >1377

rRNA  
BASE COUNT 361 a 299 c 413 g 304 t  
ORIGIN

Query Match 100.0%; Score 18; DB 1; Length 1377;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTCGGGAAACTGAGG 18  
|||||  
Db 112 GCTTCGGGAAACTGAGG 129

Search completed: June 19, 2003, 09:20:16  
Job time : 1876 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 08:25:58 ; Search time 1791 Seconds

(Without alignments)  
162.769 Million cell updates/sec

Title: US-09-911-860A-3

Perfect score: 18

Sequence: 1 gcttcggaactgaag 18

Scoring table: IDENTITY NUC

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18	100.0	248	10	BB073520 BB073520
2	17	94.4	434	12	BG812827 BG812827
3	16.4	91.1	182	13	BJ528210 BJ528210
C 4	16.4	91.1	237	9	AV288693 AV288693
C 5	16.4	91.1	243	10	BB168908 BB168908
C 6	16.4	91.1	244	10	BB201074 BB201074

C 7	16.4	91.1	276	10	BB229355 BB229355
8	16.4	91.1	626	13	BJ014233 BJ014233
9	16.4	91.1	637	13	BG994343 BG994343
10	16.4	91.1	643	13	BJ011936 BJ011936
11	16.4	91.1	691	17	BH729453 BH729453
12	16.4	91.1	698	17	A2703698 A2703698
13	16.4	91.1	633	13	A2099636 A2099636
C 14	16.4	91.1	662	13	B1971320 B1971320
15	16.4	88.9	1055	14	BM811393 BM811393
16	16.4	88.9	230	10	BB566803 BB566803
C 17	15.4	85.6	238	10	BB199049 BB199049
C 18	15.4	85.6	244	9	AV293267 AV293267
C 19	15.4	85.6	249	9	AA499732 AA499732
C 20	15.4	85.6	256	10	BB853957 BB853957
21	15.4	85.6	258	10	BE456521 BE456521
C 22	15.4	85.6	268	9	AA674970 AA674970
C 23	15.4	85.6	268	10	BB574204 BB574204
C 24	15.4	85.6	270	9	A1939806 A1939806
C 25	15.4	85.6	273	10	AV329591 AV329591
C 26	15.4	85.6	273	10	BE332265 BE332265
C 27	15.4	85.6	277	10	AV347556 AV347556
C 28	15.4	85.6	277	10	AV376069 AV376069
C 29	15.4	85.6	278	10	BB18352 BB18352
C 30	15.4	85.6	281	10	BB154538 BB154538
31	15.4	85.6	282	13	BM198714 BM198714
C 32	15.4	85.6	285	10	BB110943 BB110943
C 33	15.4	85.6	290	10	BB472639 BB472639
C 34	15.4	85.6	296	10	BB565684 BB565684
C 35	15.4	85.6	300	10	BB521050 BB521050
C 36	15.4	85.6	307	9	AV248001 AV248001
C 37	15.4	85.6	310	10	BE095654 BE095654
C 38	15.4	85.6	313	10	BB095015 BB095015
39	15.4	85.6	314	10	BB251400 BB251400
40	15.4	85.6	316	12	BF451432 BF451432
41	15.4	85.6	321	12	BG381377 BG381377
42	15.4	85.6	339	9	BB251395 BB251395
43	15.4	85.6	347	9	AA410333 AA410333
44	15.4	85.6	355	10	BB872882 BB872882
45	15.4	85.6	355	10	BB872957 BB872957

## ALIGNMENTS

## RESULT 1

BB073520/C

LOCUS BB073520 RIKEN full-length enriched, adult male epididymis Mus 248 bp mRNA EST 27-JUN-2000

DEFINITION musculus CDNA clone 9230109F19 3', mRNA sequence.

ACCESSION BB073520

VERSION BB073520.1 GI:8583518

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 248) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

## AUTHORS

Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carrinici,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayasu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Iizawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,T., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya,T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE RIKEN Mouse ESTs (Konno,H., et al.)

COMMENT Unpublished (2000) Laboratory for Genome Exploration Research Group, RIKEN Genomic







Db 101 GCTTCGGGAACCTGAGC 84

RESULT 7  
BB229355/c  
LOCUS BB229355 276 bp mRNA linear EST 03-JUL-2000  
DEFINITION BB229355 RIKEN full-length enriched, 3 days neonate thymus Mus  
muscusl cDNA clone A630020P19 3', mRNA sequence.  
BB229355  
VERSION BB229355.1 GI:8900000  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shimamoto, T., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Matabiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Kono, H., et al.)  
Unpublished (2000)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gs.c.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoke, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermosensitization and thermoinactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (3), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.  
Location/Qualifiers  
1. 276  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="A630020P19"  
/clone\_1lb="RIKEN full-length enriched, 3 days neonate thymus"  
/tissue\_type="thymus"  
/dev\_stage="3 days neonate"  
/lab\_host="DH10B"  
/note="Site 1: SalI. Site 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GACAGACAGACATCCAGAGCTCTTTTCTTTTCTTTTCTT 3'], cDNA was prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot - 20.0 and subtraction to Rot - 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GACAGACAGATTCGAGTTAAATTAAATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLX I."

BASE COUNT 65 a 54 c 42 g 115 t  
ORIGIN

Query Match 91.1%; Score 16.4; DB 10; Length 276;  
Best Local Similarity 94.4%; Pred. No. 1.1e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 128 GCTTCGGGAACCTGAGG 111

RESULT 8  
BJ014233  
LOCUS BJ014233 626 bp mRNA linear EST 05-DEC-2001  
DEFINITION BJ014233 MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA197E08 5', mRNA sequence.  
BJ014233  
VERSION BJ014233.1 GI:17361113  
KEYWORDS EST.  
SOURCE Japanese medaka.  
ORGANISM Oryzias latipes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha; Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.  
1 (bases 1 to 626)  
Kohara, Y., Shin, I., Kimura, T., Narita, T., Jindo, T. and Takeda, H.  
Medaka EST Project in Takeda's Lab  
Unpublished (2001)  
Contact: Tadasu Shin-I  
Center for Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshin@genes.nig.ac.jp.  
Location/Qualifiers  
1. 626  
/organism="Oryzias latipes"  
/strain="hd-r"  
/db\_xref="taxon:8090"  
/clone="MF01SSA197E08"  
/clone\_1lb="MF01SSA cDNA"  
/sex="mixture of female and male"  
/tissue\_type="whole embryo"  
/dev\_stage="segmentation stage 20 - 25"  
BASE COUNT 150 a 163 c 163 g 150 t  
ORIGIN

Query Match 91.1%; Score 16.4; DB 13; Length 626;  
Best Local Similarity 94.4%; Pred. No. 1.3e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 583 GCTTCGGGAACCTGAGG 600

RESULT 9  
BG994343  
LOCUS BG994343 637 bp mRNA linear EST 13-JUN-2001  
DEFINITION PMO-HT1166-130201-003-f03 HT1166 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG994343  
VERSION BG994343.1 GI:14398413

**KEYWORD** EST.  
**SOURCE** human.  
**ORGANISM** Homo sapiens  
**REFERENCE** Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.  
**AUTHORS** 1 (bases 1 to 637)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, H., Bala, G.S., Simpson, D.H., Brustein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
**TITLE** Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
**MEDLINE** 20202663  
**COMMENT** Contact: Simpson A.J.G.  
 Laboratory for Cancer Research  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM06t2-PM0-HT1166-130201-003-f03&t3=2001-02-13&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 593.

**FEATURES**  
**source**  
 1. 637  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="HT1166"  
 /dev\_stage="Adult"  
 /note="Organ: head, neck; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
 BASE COUNT  
 146 a 190 c 175 g 126 t

**ORIGIN**  
 Query Match 91.1%; Score 16.4; DB 13; Length 637;  
 Best Local Similarity 94.4%; Pred. No. 1.3e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

**QY** 1 GCTTCGGGAACCTGACAGG 18  
 |||||  
 21 GCTTGGGAACCTGACAGG 38

**Db** 21 GCTTGGGAACCTGACAGG 38

**RESULT 10**  
 B011936 643 bp mRNA linear EST 05-DEC-2001  
**LOCUS** B011936 MF01SSA cDNA Oryzias latipes cDNA MF01SSA169C01 5',  
**DEFINITION** mRNA sequence.  
**ACCESSION** B011936  
**VERSION** B011936  
**KEYWORDS** B011936.1 GI:17360330  
**SOURCE** EST.  
 Japanese medaka.  
 Oryzias latipes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
 Belontiiformes; Adrianichthyidae; Oryziatidae; Oryzias.  
**REFERENCE** 1 (bases 1 to 643)  
 Kohata, Y., Shin-I, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.  
**TITLE** Medaka EST Project in Takeda's lab  
**JOURNAL** Unpublished (2001)

**COMMENT** Contact: Tadasu Shin-I  
 Center for Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshtn@genes.nig.ac.jp.

**FEATURES**  
**source**  
 1. 643  
 /organism="Oryzias latipes"  
 /strain="HD-r"  
 /db\_xref="taxon:8090"  
 /clone="MF01SSA169C01"  
 /clone\_lib="MF01SSA cDNA"  
 /sex="mixture of female and male"  
 /tissue\_type="whole embryo"  
 /dev\_stage="segmentation stage 20 - 25"  
 BASE COUNT  
 158 a 166 c 166 g 153 t

**ORIGIN**  
 Query Match 91.1%; Score 16.4; DB 13; Length 643;  
 Best Local Similarity 94.4%; Pred. No. 1.3e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

**QY** 1 GCTTCGGGAACCTGACAGG 18  
 |||||  
 606 GCTTCGGGAACCTGACAGG 623

**Db** 606 GCTTCGGGAACCTGACAGG 623

**RESULT 11**  
 BH729453 691 bp DNA linear GSS 20-FEB-2002  
**LOCUS** BH729453  
**DEFINITION** BOMGB67Tf BO\_2\_3\_KB Brassica oleracea genomic clone BOMGB67, DNA  
 sequence.  
**ACCESSION** BH729453  
**VERSION** BH729453.1 GI:18834848  
**KEYWORDS** GSS.  
**SOURCE** Brassica oleracea  
 Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.  
**REFERENCE** 1 (bases 1 to 691)  
 Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.  
**AUTHORS** Whole genome shotgun sequencing of Brassica oleracea  
**TITLE** Unpublished (2001)  
**JOURNAL** Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TF  
 Class: sheared ends.

**FEATURES**  
**source**  
 1. 691  
 /organism="Brassica oleracea"  
 /strain="T01000DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOMGB67"  
 /clone\_lib="BO\_2\_3\_KB"  
 /note="Vector: PHOS1; Site\_1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers."  
 BASE COUNT  
 161 a 165 c 153 g 212 t

**ORIGIN**  
 Query Match 91.1%; Score 16.4; DB 17; Length 691;  
 Best Local Similarity 94.4%; Pred. No. 1.3e+03;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

**QY** 1 GCTTCGGGAACCTGACAGG 18  
 |||||



Db 608 GCTTCGGGAACCTGAAG 625

RESULT 12  
LOCUS A209636 698 bp DNA linear GSS 24-JAN-2001  
DEFINITION RPCI-23-225D9.TJ RPCI-23 Mus musculus genomic clone RPCI-23-225D9,  
DNA sequence.  
ACCESSION A209636  
VERSION A209636.1 GI:12428344  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
1 (bases 1 to 698) Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinet  
and Fraser,C.M., B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.  
and Fraser,C.M.  
TITLE Mouse BAC End Sequences from Library RPCI-23  
COMMENT Unpublished (1999)  
Other\_GSSs: RPCI-23-225D9.TJ  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhaoc@tigr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC  
library availability, please contact Pieter de Jong  
(pdejong@tigr.org). Clones may be purchased from BACPAC  
Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end  
plate: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 225 row: D column: 9  
Seq primer: SP6  
Class: BAC ends.  
Location/Qualifiers  
1..698  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-225D9"  
/clone\_1lb="RPCI-23"  
/sex="Female"  
/lab\_host="DH10B"  
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site\_1:  
ECORI; Site\_2: ECORI; Female C57BL/6J mouse kidney and/or  
brain genomic DNA was isolated and partially digested  
with a combination of EORI and EORI Methyase. Size  
selected DNA was cloned into the pBAC3.6 vector at the  
ECORI sites. The ligation products were transformed into  
DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 209 a 131 c 147 g 211 t

ORIGIN

Query Match 91.1%; Score 16.4; DB 17; Length 698;  
Best Local Similarity 94.4%; Pred. No. 1.3e+03;  
Matches 17: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACCTGAAG 18  
|||||  
Db 671 GCTTCGGGAACCTGAAG 688

RESULT 13  
LOCUS A209636 633 bp DNA linear GSS 09-MAY-2000  
DEFINITION RPCI-23-464K4.TJ RPCI-23 Mus musculus genomic clone RPCI-23-464K4,  
DNA sequence.  
ACCESSION A209636  
VERSION A209636.1 GI:7752692  
KEYWORDS GSS.

SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
1 (bases 1 to 633) Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinet  
and Fraser,C.M., B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.  
and Fraser,C.M.  
TITLE Mouse BAC End Sequences from Library RPCI-23  
COMMENT Unpublished (1999)  
Other\_GSSs: RPCI-23-464K4.TJ  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhaoc@tigr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC  
library availability, please contact Pieter de Jong  
(pdejong@tigr.org). Clones may be purchased from  
BACPAC Resources (<http://bacpac.med.bufile.edu/orderingframe.htm>)  
or from Resea ch Genetics ([info@resgen.com](http://resgen.com)). BAC end page:  
[http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 464 row: K column: 4  
Seq primer: SP6  
Class: BAC ends.  
Location/Qualifiers  
1..633  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-464K4"  
/clone\_1lb="RPCI-23"  
/sex="Female"  
/lab\_host="DH10B"  
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site\_1:  
ECORI; Site\_2: ECORI; Female C57BL/6J mouse kidney and/or  
brain genomic DNA was isolated and partially digested  
with a combination of EORI and EORI Methyase. Size  
selected DNA was cloned into the pBAC3.6 vector at the  
ECORI sites. The ligation products were transformed into  
DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 143 a 176 c 161 g 153 t

ORIGIN

Query Match 88.9%; Score 16; DB 17; Length 633;  
Best Local Similarity 100.0%; Pred. No. 2e+03;  
Matches 16: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTTCGGGAACCTGAA 16  
|||||  
Db 222 GCTTCGGGAACCTGAA 237

RESULT 14  
LOCUS B1971320/c 662 bp mRNA linear EST 23-OCT-2001  
DEFINITION GM830013A11E09 Gm-r1083 Glycine max cDNA clone Gm-r1083-4817 3',  
mRNA sequence.  
ACCESSION B1971320  
VERSION B1971320.1 GI:16345725  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
REFERENCE 1 (bases 1 to 662) Vodka,L., Keim,P., Shomaker,R., Retzel,E., Khanna,A., Coryell,V.,  
Erpelding,J., Raph,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.  
A Functional Genomics Program for Soybean (NSF 9872565)